DOCUMENT RESUME

ED 299 926 HE 021 925

TITLE Contexts for Assessment. AAHE Bulletin. October 1988.
INSTITUTION American Association for Higher Education, Washington, D.C.
PUB DATE Oct 88
NOTE 15p.
AVAILABLE FROM American Association for Higher Education, One Dupont Circle, Suite 600, Washington, DC 20036 ($3.00).
PUB TYPE Collected Works - Serials (022)
JOURNAL CIT AAHE Bulletin; v41 n2 p1-14 Oct 1988

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Planning; Higher Education; Role of Education; School Effectiveness; Student Evaluation; Writing Evaluation; *Writing Skills
IDENTIFIERS *College Outcomes Assessment; Educational Issues

ABSTRACT

Three articles on assessment in higher education include the following: "More Savvy, Good Signs, Next Steps: A Report on the Third National Conference on Assessment in Higher Education," by Pat Hutchings, on tapping expert judgment, getting students invested, and deepening understanding; "Perspectives on General Education and Its Assessment," by Ted Marchese, an interview with John Harris and Jerry Gaff; and "Process vs. Product: Assessing Skills in Writing," by Edward M. White. (SM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

OCTOBER 1988 VOLUME 41 NUMBER 2

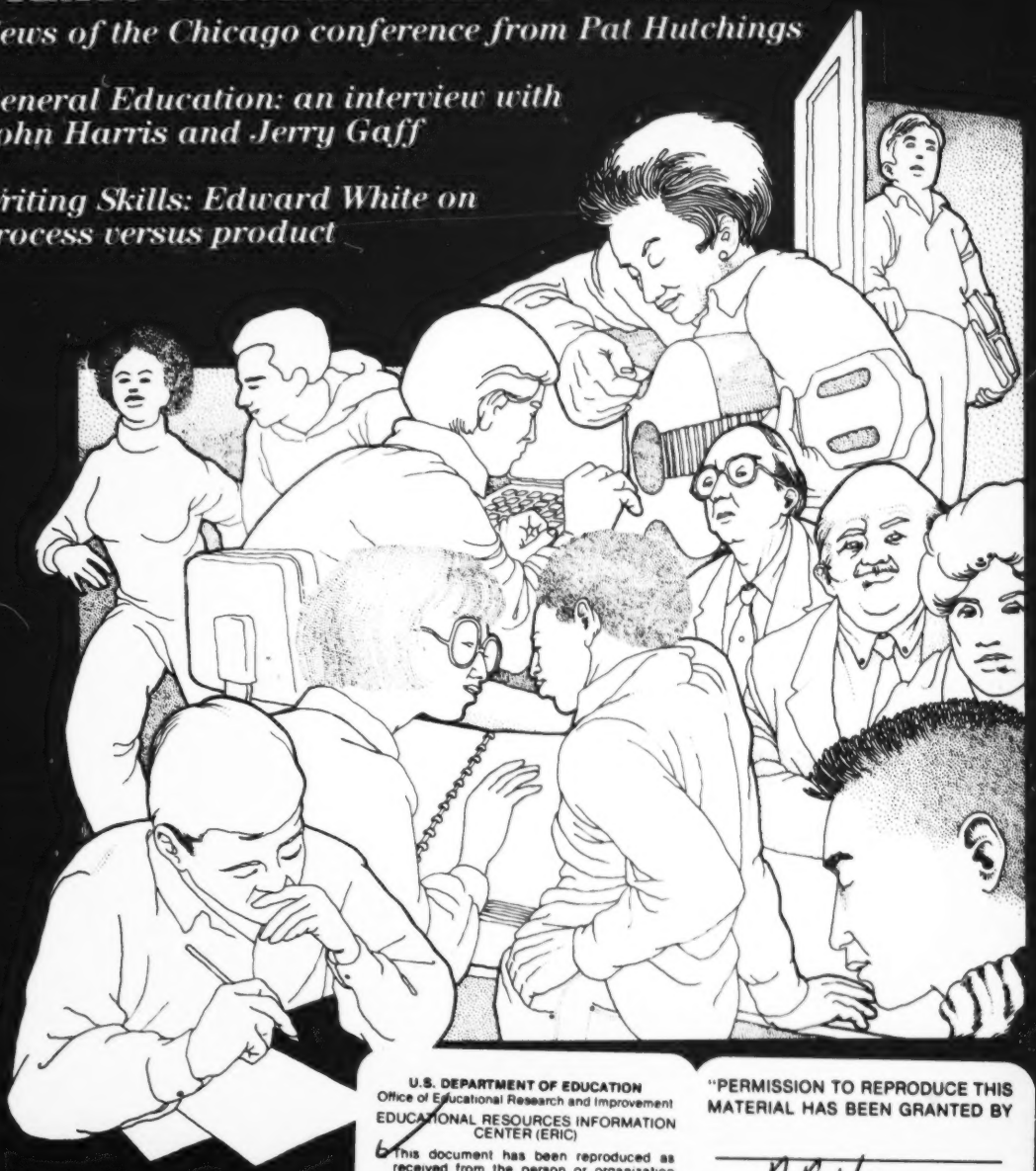
BULLETIN

AAHE
AMERICAN ASSOCIATION
FOR HIGHER EDUCATION

ED299926

CONTEXTS FOR ASSESSMENT

- *News of the Chicago conference from Pat Hutchings*
- *General Education: an interview with John Harris and Jerry Gaff*
- *Writing Skills: Edward White on process versus product*



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
AAHE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

HE021925T

B U L L E T I N

OCTOBER/1988 VOLUME 41/NUMBER 2

In This Issue:

More Savvy, Good Signs, Next Steps — *by Pat Hutchings* — 3
A report on the Third National Conference on Assessment in Higher Education

Perspectives on General Education and Its Assessment — *an interview with John Harris and Jerry Gaff* — 6

Process vs. Product — *by Edward M. White* — 10
Assessing skills in writing

Bulletin Board — *by Ted Marchese* — 15

News & Information — 16

Peter Ewell, writing in the July/August *Change*, describes the danger in state policy of assessment becoming "a train on its own track," an initiative, that is, separated from longer-run mechanisms for budgeting, planning, or program review. Thus separated, assessment develops its own "paper trails and bureaucratic paraphernalia" but not the longer-term linkages that give it impact and staying power.

Fortunately, as we saw this past June at AAHE's assessment conference in Chicago, most campuses seem well aware of that danger and aim to make the assessment they do integral to programs of undergraduate improvement. An oft-cited analogy, of assessment to a bathroom scale, captures that understanding: a scale does not cause you to lose weight. Dieting and exercise do. But in a *program* to lose weight, the scale plays an essential role: it provides feedback about results.

Two points follow from the analogy. One is that what's being discovered on campuses about assessing student learning needs the attention of the larger, diverse communities pursuing curricular and instructional reform; they need to hook the assessor's question and mechanisms—What, specifically, is the intended result? How can we know of its achievement?—into their prescriptions. A second is that the assessor's work properly begins not with questions of technique but with a sense of the values and issues behind a larger program of improvement.

These perspectives undergirded our Chicago conference and run through Pat Hutchings's report in this issue on the event. What is it, she asks, that we value in faculty? For students? Let those values direct what is done in the name of assessment.

Two other pieces in this *Bulletin* derive from Chicago presentations. Ed White's essay makes the point that how you assess student writing depends first on how you think about the teaching and learning of writing. Similarly, as the Gaff-Harris piece indicates, the place to start in assessing general education outcomes is not with matters of technique but with your intents for general education.

Assessment is interesting (and potentially powerful) because done right it drives you back to first questions of values and purpose. Because those questions are AAHE's questions, and because assessment, in Jerry Gaff's words, is "where the updraft is," we'll stay with them in future *Bulletins*.

—TJM

AAHE BULLETIN

October 1988, Volume 41, Number 2

Published by
American Association for Higher Education
One Dupont Circle, Suite 600, Washington, DC 20036
Telephone: (202) 293-6440
President: Russell Edgerton
Vice President: Theodore J. Marchese
Managing Editor: Bry Pollack

AAHE Bulletin (ISSN 0162-7910) is published as a membership service of the American Association for Higher Education, a nonprofit organization incorporated in the District of Columbia. Second class postage paid at Washington, DC. Annual domestic membership dues: $65, of which $35 is for publications. Subscription price for *AAHE Bulletin* without membership: $27 per year, $35 outside the U.S. *AAHE Bulletin* is published 10 issues per year, monthly except July and August. Back issues: $3 each for up to 10 copies; $2 each for 11 or more copies. Payment must accompany all orders under $25; purchase order must accompany all orders over $25. **POSTMASTER**: Send address changes to *AAHE Bulletin*, One Dupont Circle, Suite 600, Washington, DC 20036.

Typesetting by Ten Point Type
Printing by Hagerstown Bookbinding and Printing, Inc.

Cover art and design by Max Karl Winkler

More Savvy, Good Signs, Next Steps

A report on the Third National Conference on Assessment in Higher Education

by Pat Hutchings

Much of the assessment action over the past few years has been *re*action—to state mandates, to new accreditation criteria, and to repeated calls for higher education to prove itself to a skeptical public. At the FIPSE-funded Assessment Forum's Third National Conference on Assessment in Higher Education, a shift in the wind was evident.

Educators are taking hold of assessment, making it their own, shaping it to serve the purposes of teaching and learning.

Like our previous assessment gatherings (Columbia, SC, in 1985 and Denver in 1987), this year's event filled the house. Not surprisingly, public institutions were in the majority, but private colleges and universities were a significant presence as well, constituting a third of the total. Many participants attended as campus teams, and faculty enrollment was up from last year.

All in all, more than 1,000 people gathered at the Palmer House in Chicago, June 8–11, for four days of workshops, panels, team meetings, and plenary addresses. With rooms still filled to capacity for 4 PM sessions, it was clear this crowd had come to work.

In fact, there was more this year to work *at*. The program reflected deepening experience, broader options, more savvy among presenters and audience alike. There was plenty of skepticism to be heard, and appropriately so, but I was struck by several more promising signs.

Three years ago, when assessment was first gaining momentum, there was widespread concern that it might trigger mainly quick-fix, compliance activities. Not so, if the Chicago group is any indication.

Consider, for instance, the open forum on portfolio review late Thursday afternoon. Portfolios present rich possibilities for understanding the student experience, but they can also be unwieldy, time-consuming, and labor-intensive. My hope was that ten or fifteen brave souls would be intrigued enough to come talk about such methods; instead, the session drew more than a hundred people.

The use of portfolios to assess a wide range of student knowledge and abilities (not only writing) is a strategy for further work—in follow-up sessions at our 1989 conference, for instance—but more on that in a bit.

Then there's self-assessment. For those who see assessment first and foremost as a way of prompting individual student learning, self-assessment is key. It is, if you will, a way to turn "students" into "learners," and several institutions—Alverno College, for one—actually assess their students' ability to self-assess.

Pat Hutchings is the director of AAHE's Assessment Forum.

Like portfolio review, a focus on self-assessment requires significant commitments of time and energy, and it's not the sort of method state legislators are likely to reward. But when the School for New Learning at De Paul University offered a two-hour workshop on using self-assessment to empower student learning, it drew a crowd three times the anticipated size. . . . Another issue to explore further next year.

The point, in both of these examples, is that educators are steering assessment in educational directions. This is not to say that state-level assessment initiatives don't continue to shape the scene—or that they don't have a legitimate role. It is, rather, to highlight growing interest in methods that give a fuller, more complete picture of which students learn what under what conditions.

Rather than single score "snapshots," we're seeing more and more qualitative, performance-based assessment.

In addition to portfolio review and self-assessment strategies, the Chicago sessions looked at the use of focus groups (Virginia Commonwealth University uses them to assess student values) and interviews (used to assess general education outcomes at a consortium of nine Appalachian colleges). Using assessment information for better advising was a repeated Chicago theme, as were assessment center methods, expert judgment, assess-

ment in the classroom, "hard-to-assess" abilities such as response to the arts. . . .

These are encouraging signs. Rather than sitting back watching mandates close in, nooselike, many campuses are taking assessment matters into their own hands, finding ways to do things that make sense for them and for students. The question becomes, what do we need to do to keep on this promising road?

There's no eight-step route to happiness, but the experience of many campuses suggests two broad lessons to keep in view—one about faculty, the other about students. Neither is new but both are worth repeating and, with the greater experience now available, elaborating.

Tapping expert judgment

In the thick of forming committees, designing plans, and choosing instruments, it's easy to lose sight of the larger, long-standing fact that faculty are the single most powerful assessment resource on campus. Faculty expert judgment, if tapped appropriately, can yield a wealth of information about what and how students learn.

Entry-level assessment brings this point home. Energetic R&D from the testing companies (much more work than has gone into general education, for instance, where options are still thin) has given us a number of useful entry-level instruments—most recently, computer-adaptive methods that are highly diagnostic. But even the "perfect" test, if there were such a thing, would give only a small piece of the total entry-level picture we need.

What's really at stake at entry, after all, is effective placement and advisement, and for that one needs an array of information: a test score, yes, but also information on student background, goals, work, financial needs, previous studies and grades, learning style, self-knowledge. Most of all, we want the judgment of faculty brought to bear in sorting out such information and using it to interpret and predict actual student work.

For example, under what conditions will this student do her best work? Will she perform more effectively with coaching and encouragement, or in a more challenging, high-pressure setting such as placement one level above what a test score might indicate? Faculty experience should weigh far more than a test score in making such judgments.

The power of faculty expert judgment is also apparent in assessment in the major, an area that attracted standing-room-only attention in Chicago.

As faculty grapple with learning in their disciplines, the need for more context-sensitive assessment methods becomes apparent. A presentation by Kean College faculty from six different areas illustrated that different disciplines call for different methods, that the meaning of performance and expert judgment of performance varies by subject area.

An implication here, as in entry-level assessment, is that testing can never be enough. The Educational Testing Service is currently scaling down the GRE area tests to make them more suitable for departmental assessment and program evaluation; it's a development that warrants attention.

Particularly in the case of upper-level learning in the major, however, the need is for methods that call forth the full range of what students know and can do. There's a need for approaches that get at the more subtle outcomes that faculty care about, the sensibilities and dispositions that distinguish, say, the English major. There's a trap in thinking everything can and should be measured, but surely faculty have ways to get a reading on such slippery outcomes as, in English, a care about language. One might, for starters, look more carefully at how students revise their own writing.

One dimension of assessment—and one of its strengths—is a call for externality, for someone besides the cook to taste the pudding. But this recipe need not exclude a faculty's long-standing experience with particular students encountering particular subject matter.

Indeed, one way to achieve externality is to make more public what faculty know—and what they want to know—about student learning, about standards and judgments, about the connections between learning in courses and other contexts. . . . The point, quite simply, is that powerful assessment begins with the question Kean College faculty used as the title for their session: "What do faculty want to know?"

Getting students invested

Much has been said, in the name of assessment, about accountability—said mainly by those outside academe. As pointed out by a faculty member featured in a recent *Change* article on faculty views of assessment, however, "our ultimate accountability" is to students.

It's a line that bears repeating, since one of the concerns that surfaced repeatedly in Chicago is the difficulty of getting students invested in assessment. It seems they've gotten wind that "these tests don't count," that there's nothing at stake. The result (free pizzas not withstanding) is students who won't take the tests or, even more problematic, won't take them seriously.

A number of institutions attempt to involve students in assessment plans—with advisory councils, articles in the school newspaper, and so on. But assessment needs to give further attention to students and how they experience assessment.

For starters, it's time to recognize that our choices of assessment methods and instruments send powerful messages to students. Those messages were a central theme in Alexander Astin's presentation, "Assessment and Human Values." Speaking as a "reformed number cruncher," Astin described an "ethic of competition" underlying norm-referenced standardized tests—where quality is defined as doing better than the next person.

That's not to say that such tests have no place in assessing student learning outcomes, or that comparative data is not occasionally illuminating. It is to say that we need to have in mind considerations beyond the technical and administrative when we choose our methods. What we measure, and how, sends a message—and students get it—about what the institution values, about what matters.

Moreover, the methods we

choose convey attitudes about learning and about our expectations of students. If we ask students to take tests that call for right answers selected from among predetermined alternatives (multiple-choice), what, Astin asked, do we thereby communicate about the need for creativity? About the need to apply learning to solve the kind of open-ended problems we confront in life? What does it mean that so few campuses assess (and therefore endorse) the ability to listen or to cooperate?

There's a related issue here as well, and it takes us back around to the Chicago session on self-assessment that I mentioned earlier. Assessment sends messages by virtue of its *what* and *how*, but there's also an issue of *who*. Research—from the Learning and Retention Consortium (LARC) of California community colleges, for instance—shows that students, at entry, overestimate their skill as writers; they are unreliable assessors of their own work.

Probably so. But the ability to judge performance is an acquirable one, and the importance of doing so is at the heart of assessment. If it's important for institutions to examine their effectiveness and for faculty to evaluate teaching, then surely this is a habit of mind we want to teach students as well.

It's interesting to think what would happen if we put as much energy and money into teaching students to make judgments about their own learning as we do into tests.

Deepening understanding

In last September's *Bulletin*, reflecting on the 1987 Denver assessment conference, I reported a theme of "first questions" and the power of assessment to take us back to issues of quality and purpose. Such issues were evident this year, too. What was new in Chicago is a deepening understanding of them in light of faculty and student experiences of assessment.

It's a focus that bodes well for the next year of assessment—and for next year's conference, the Fourth National Conference on Assessment in Higher Education, June 21-24 in Atlanta. Hope to see you there. ■

MARK YOUR CALENDARS!

The Fourth National Conference on Assessment in Higher Education

June 21-24, 1989 ■ The Atlanta Hilton and Towers ■ Atlanta, Georgia

Last June's Assessment Forum conference in Chicago brought together more than 1,000 representatives from campuses across the country, ready to move ahead with assessment. In Chicago, participants discovered thoughtful speakers, practical advice, rich resources, and new colleagues. The next step is Atlanta, where we invite you to join us June 21-24 at the Atlanta Hilton and Towers.

Highlights of the 1989 Assessment Forum conference include:

- On-going, in-depth tracks on areas including assessment in the major, assessment in community colleges, and assessment of general education.
- Sessions on a broad range of assessment options, from entry-level to alumni follow-up.
- Updates on the latest tools and instruments.
- A look at state policy on assessment and strategies for responding.
- Campus teams—a chance to work together and enjoy discounted registration fees.
- A preconference round of limited-enrollment workshops.
- Opportunities for post-conference meetings by campus and state delegations.

For further details and registration materials, contact:

THE AAHE ASSESSMENT FORUM

One Dupont Circle, Suite 600
Washington, DC 20036
(202) 293-6440

Perspectives on GENERAL EDUCATION AND ITS ASSESSMENT

an interview with John Harris and Jerry Gaff

John Harris is a faculty member at David Lipscomb College (TN), with a background in instructional design and educational evaluation. In the 1970s, Harris contributed to the literatures on competency-based education and outcomes analysis. In 1985, he wrote a widely cited paper on assessing general education outcomes.

Since then, he has been a sought-after expert in assessing general education and has consulted on numerous campuses.

Jerry Gaff brings to assessment a background in psychology. In the 1970s he became known for his role in two national studies: one of faculty development in higher education, the second a survey of practices in general education, which led to his book General Education Today.

For the past five years, Gaff has been dean and acting president at Hamline University (MN), at which he helped develop a new general education curriculum and a plan for assessing it.

AAHE vice president Ted Marchese interviewed Harris and Gaff in a session at the Third National Conference on Assessment in Chicago this past June.

Harris

Gaff

MARCHESE: The assessment of general education outcomes has attracted considerable attention lately, among accreditors, faculty committees, and state authorities. Why all this interest? And, why is this question worth our time? Jerry?

GAFF: I have two answers. First, there's interest because general

education is a "disaster area," and we need to make repairs. That was what the Carnegie Foundation for the Advancement of Teaching called it back in 1977, in a report that helped touch off a whole reform movement in general education. That "disaster" label still applies on a large number of campuses.

The second reason is that general education is *not* a disaster area. Many campuses have indeed instituted significant general education reforms this decade. Now they need to find out how those reforms are working and what further changes should be made. Assessment is a tool for answering those questions.

MARCHESE: John, you've been on a lot of campuses, private ones especially. What do you see behind their interest in assessing general education?

HARRIS: For some of the smaller private colleges, it's an issue in the air, and they feel they should be addressing it—mandate or not. The question then is, why is it in the air?

In the last ten years our culture has, I think, become increasingly conscious of quality as an issue. In the world of higher education, we have a fair amount of evidence now that more Ph.D.'s and more books in the library—the traditional indexes of quality—don't necessarily assure quality of result. So there's a call for more direct measures of quality, to look at the outcomes of programs and instruction, and to ask what and how students have learned. Those places that are sincerely interested in making systematic, long-term improvements know that there are no magic answers out there, assessment or anything else, but that the only way to start on the path to improvement is by tracking the results of our efforts.

MARCHESE: Before we go any further, I want to tackle the definition problem. The Association of American Colleges (AAC) recently issued a report, *A New Vitality in General Education,* that defines general education as "the cultivation of the knowledge, skills, and attitudes that all of us use and live by for most of our lives, whether as parents, citizens, lovers, travelers, participants in the arts. . . ." That sounds like a definition of *all* of higher education. What are we talking about when we say "general education"? Does the phrase have any concrete meaning that can drive a conversation about assessing it?

> While educators are talking about lofty goals, the public is concerned first about the more basic abilities of our graduates.

GAFF: The answer ought to be yes, but it isn't. Art Levine and Ernie Boyer wrote a book called *The Quest for Common Learning,* and in an appendix they listed twenty-five or thirty definitions that have been given for general education.

What it comes down to is that general education is whatever a campus decides it is, when that campus tackles questions about what students need to know and be able to do in order to lead productive, fulfilling lives.

There are some common answers to that question, but a lot of differences, too. The AAC definition, which comes from Joe Katz, is probably as good a one as you're likely to come up with, but the fact remains that the operational definition is a social one that comes out of actual campus practice.

HARRIS: I define general education as whatever a faculty believes ought to be common among all of its graduates, regardless of the students' major or career.

MARCHESE: Most generic definitions of general education entail talk about broader, degree-level outcomes.

HARRIS: Yes, and I'd like to advance a thesis: Whatever the broad, global outcomes of an institution will be—if nothing else, they will be very difficult to agree upon—it's particularly important to agree on some common, higher literacy and—to use the British term—"numeracy" skills that students cannot graduate without.

GAFF: In most institutions, general education points to some common minimal expectations, but also to such things as understanding different disciplines and their ways of knowing, integrating or drawing conclusions across disciplines, understanding global contexts, studying foreign languages. . . . There are a lot of very important and legitimately different definitions. The trick, on any campus, is to weave together those different meanings. Each meaning comes out of a particular political and social context, and a big part of building a general education program is building political and social support. Defining general education necessarily involves choices, and that means compromises.

HARRIS: The public perspective is relevant here. I think the public—including groups that are mandating assessment of general education—is very willing to have us pursue larger outcomes. But while educators are talking about lofty goals, the public is concerned first about the more basic abilities of our graduates.

MARCHESE: The oft-cited gospel is that you can't do assessment unless you have clarity about goals . . . in this case, about what you expect students to know and be able to do upon award of the degree. Given the variability and mushiness of general education goals, can it make sense to assess the function? Is the place to start with the goals?

HARRIS: Ideally, yes. But what one finds in reality is that only colleges with high social cohesion are likely to be able to state specific enough goals. Faculty in larger, more complex institutions tend to be diverse

and atomistic in their thinking and are less likely to be able to formulate concrete overall goals.

What I suggest for most campuses, then, is that they begin with empirical rather than theoretical questions about general education. Find out what experiences now are common among students. Look at the courses most frequently taken by all students, at the examinations and requirements, to see what content is covered, what level of performance is being required, and so forth.

It's sometimes a startling exercise. Faculty say they're teaching higher-level skills, but when you look at their examinations, they're testing for regurgitation of knowledge.

MARCHESE: Is the point of such inquiries to come up with a goal statement?

HARRIS: My suggestion is that while the logic is to begin with a full set of clearly stated goals, assessment may have to start with empirical questions and a partial set of goals, with questions about those things that *can* be agreed upon—for example, the importance of writing.

I've found that often the most important goals or commitments on a campus are not that easy to state clearly. So it's important to find out what common values and commitments exist on the campus.

I think that Jerry and I agree that general education is an expression of the collective educational values of the campus. There may be some collective values or goals there that are not immediately assessable. I don't mean to jump ahead, but I'd say that if you can't completely assess a goal, like with a test or a student performance, you should look for at least some rough indicators of the goal's affects, and start with those.

GAFF: Ted, your questions about goals are crucial to assessment. Most of the problems we encounter with assessing general education are actually problems with general education itself. What we need—and don't really have in many cases—are goals that are specific enough to drive courses and outcomes analyses but broad enough to allow flexibility for individual faculty, courses, and students.

> We need goals that are specific enough to drive courses but broad enough to allow flexibility for faculty and students.

MARCHESE: What you're saying, Jerry, is that it's important not to let the psychometricians take over assessment by insisting on a kind of goal statement that, in striving toward measurable, behavioral terms, ignores the broader purpose of goals, which is to advance education, not evaluation.

GAFF: Absolutely.

MARCHESE: Most colleges, while they have something they call general education, don't have a *program* of common coursework that all students take—which makes it quite difficult to focus the assessment on content. That drives you back to skills and abilities. Do you have any uneasiness about assessing skills and abilities independent of their exercise in particular knowledge areas?

HARRIS: Yes, I'm concerned. The consensus among people who worry about critical thinking, to cite just one example, is that it's hard to judge that ability independent of content knowledge. Where we could come nearest to assessing the higher-order skills that Jerry talks about is in a pet area of mine: the senior thesis or project. A capstone exercise can reflect not only knowledge and skills from the major but more broad-based abilities to do critical thinking and problem solving. That's a place I'd drive a major stake in assessing general education outcomes.

GAFF: Another way is to go back to Katz's definition—to take some common issues that confront people in daily life, and ask students to apply their learning to those issues. If students aren't learning knowledge and abilities that are relevant to effective functioning in daily life, what does general education mean?

MARCHESE: What I like about both suggestions is that they lead you to the best kind of assessment: one that's a good educational experience for students.

From my own experience, there's something in the general education literature that rings not quite true. It says, begin with clear goals, apply your instruments, make the indicated changes, and so forth. It makes sense in theory, but the world seldom works that way. There are other approaches—like doing things that make sense to or for students, and gathering small bits of information to start a conversation that wasn't there before.

HARRIS: I believe that's the most important thing I might convey. Rather than spending a long time talking about general education goals, and reflecting on the theoretical and psychometric appropriateness of various approaches, I would start by *doing* something, no matter how modest.

The second thing is that whether or not you have a commitment to common goals, it's possible to have a commitment to continuous improvement, which means using whatever information is at hand, making one improvement, and then turning to a next batch of information and a next improvement. The commitment to improvement is probably the most important thing.

MARCHESE: A thing we've seen at the Assessment Forum is that people believe that assessment will be a magic bullet. If we have this information, presto, people will know what to do with it and improvements will happen. That is often far from the case. Isn't it possible that, in the name of assessment, we'll

generate a lot of information that nobody needs or wants?

GAFF: Yes, and that takes us to the issue of the governance of general education. When we started the project on general education models that I directed several years ago, we asked participants what was wrong with general education on their campus. One of the responses was "fragmentation"—that there's no single body or person responsible for planning, administering, or assessing it. A frequent outcome of the reform movement has been a clearer center, a proper governance system. With that in place, you're much more likely to have a client whose questions can drive assessment and who will be invested in using its results.

HARRIS: I was recently on an Iowa campus that's using the ACT-COMP in a longitudinal study. If the study indicates no gain, I asked them, what would that tell them about their general education program, and what would they be likely to do about it? The answer was, "absolutely nothing."

First of all, their faculty had never bought into the COMP as a measure of what they believe about general education. In the face of no gain, the faculty will simply conclude that the test isn't measuring what they're teaching. Second, what you and Jerry have been saying is true, Ted. There's no single person or group that feels accountable for general education. It's clear who's responsible for chemistry and sociology, but nobody feels responsible for the more integrated abilities and skills.

> My thesis is that to the extent that the rest of the institution supports general education, it will be effective.

GAFF: John's observation falls in the category of "things they never told us about general education." Focusing exclusively on the outcomes we want students to achieve, we forget that an institution has a whole set of arrangements that keep general education running. My thesis is that to the extent that the rest of the institution supports general education, it will be effective, and to the extent that the institution doesn't, general education is in trouble. So that when we start assessing general education, we need to look at things like admissions and what we communicate to prospective students. We need to look at advising and how faculty advisors communicate general education goals to current students. We need to look at retention, at what faculty are rewarded for, and at what gets funded at what level.

A lot of data relevant to assessing the effectiveness of general education already exists in most institutions. Elaborate new instruments may not be necessary. There's a host of easier questions that would reveal how the whole apparatus of the college either supports or undermines general education.

MARCHESE: That brings me to a last question. If improved, more powerful forms of general education are what you're after, is assessment the best tool?

HARRIS: Assessment is the right place to begin if you are seriously interested in consistent improvement of the program. But it's a long-term commitment that can't proceed without high-level leadership. In many cases, campus assessment efforts get bogged down in questions about techniques and instruments, questions that could be solved rather easily with clearer leadership about the mission of the college and where it's headed.

GAFF: The purpose of assessment is to improve education. You can go to work on improving general education in a whole host of ways, assessment being just one. To be effective, assessment needs to be part of a larger improvement effort that also includes attention to curriculum, sharper goals, course review, faculty development, and work on pedagogy. Right now assessment is where the updraft is, so I say take advantage of it. ■

DIRECTIONS FOR CONTRIBUTORS

Interested in submitting an article for publication in *AAHE Bulletin*? The following guidelines should help:

- Submissions for publication should be 1,500 to 3,000 words in length.
- All manuscripts must be typed, double-spaced, on 8½"x11" paper (one side only), with at least one inch margins.
- Manuscripts are usually published within 6 months of acceptance.
- The *Bulletin* is not refereed.
- Fresh, factual information is what we need, but extensive footnotes and references are not required.
- AAHE reserves the right to edit, shorten, or lengthen articles at the editors' discretion.
- Questions should be directed to the managing editor.

PROCESS VS. PRODUCT

Assessing skills in writing.

by Edward M. White

What do writing teachers and researchers mean when they say that writing is a* process? *What are the implications of that meaning for the measurement of writing ability?

When writing researchers, textbook writers, and teacher trainers talk about writing instruction these days, someone is sure to offer up the latest catch phrase: "Writing is a *process*, not a *product*." This slogan, like all slogans, embodies an entire complex of attitudes and approaches that bear examination.

What do writing teachers and researchers mean when they say that writing is a *process*? What are the implications of that meaning for the measurement of writing ability? And, what about the relationship of measurement to teaching, and the ways changes in teaching must affect concepts of measurement?

In the first place, it is important to notice that the last generation has seen a major change in writing research and writing instruction. Numbers of scholars, describing this change, have turned to Thomas Kuhn's metaphor for scientific revolutions and called it a "paradigm shift." Others have argued that the change is so based in the rhetorical tradition that such elaborate terminology is unseemly. The process/product distinction surely overstates the opposition that is involved in the change, but it puts the issue in stark relief.

Briefly put, this change has shifted the teaching of writing into the same context as the teaching of critical thinking and problem solving, and it has somewhat diminished the role of writing as an agent of social stratification.

I won't take space to trace the causes of this change, but I will mention a few contributing factors: the onset of open enrollment at the City University of New York in 1970, research in socio-linguistics and dialectology demonstrating the value of variant dialects, heightened concern for racial bias in instruction and measurement as part of the drive toward a more racially open society, increasing worries about the place of creative thought in a weakened school curriculum, fuller understanding by writing researchers of how writers actually produce texts, and so on. All of these issues have asked us to define writing as more than a mere product to be graded for correctness.

Of course, the slogan oversimplifies the issue. Every sensible person knows that writing is, or can be, both a process *and* a product. The current objection to defining writing as a product too easily ignores that a term paper, a letter of inquiry, a job application, an essay on a test, or an article such as this one *is* a product offered for evaluation and consumption. But

the objection is nonetheless sound, for almost all school writing is inherently valueless or meaningless as product.

We do not assign writing to students because they will turn out work that will inform us or anyone else; even doting parents, treasuring diaries and letters, have no use for term papers. We assign writing to students so that they will learn something valuable while writing. That is, school writing is necessary for instruction because it is an important—probably the single most important—way to learn.

Now, I am not telling you anything new by saying this—everyone who writes or teaches knows that writing not only communicates knowledge but also creates knowledge. But the implications for teachers (and thus for tests) are quite new.

The teacher who at heart considers school writing as a product will serve basically as a more-or-less-negative evaluator, red-marking perceived violations of some set of rules (generally, the evidence tells us, to little or no effect). Writing, when defined as a product, demands a high level of careful editing for correctness and neatness, traits that (according to fairly accurate student belief) count more than such risky matters as originality, a personal voice, or even a developed idea. According to the product model, teachers and tests serve as society's gatekeepers, screening out those students whose written products do not measure up to the standards of school correctness and the school dialect.

The quaint faith in rules for writing—the belief that if one is sufficiently correct, one will write well—not only is pervasive among old-fashioned teachers and test makers but also is firmly held by most students. Survey any group of college students and you will find that in their heart of hearts they believe that there are a handful of rules (which could be learned by sufficient drill) that when followed can transform ordinary students into good writers. If they could only learn these rules, the students believe, they could then turn out perfect first drafts, as they believe good writers invariably do.

The students, of course, don't know what these rules are; English teachers are supposed to have the inside information, but they aren't telling. So the students must guess. In fact, every student has picked up and clings to one or two of these supposed rules: Never use the word *I*; avoid all contractions; use lots of Latin terms such as *ibid*; always use the comma after the name of a tree; and so on.

Furthermore, they imagine that their fifth-grade teacher handed out the rules on the one day they were sick; thus, everyone else knows all the rules. That the supposed rules are wrong or ridiculous in no way diminishes the fervor with which the students hold them, and the students' belief that a few more rules will make them good writers never dies.

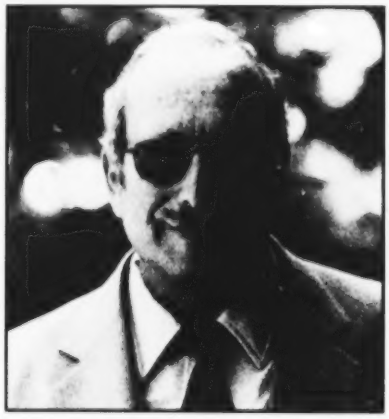
I trace this painful perplex—identifying good writing with supposed rules of correctness—to all the misplaced "grammar" instruction and multiple-choice usage testing that students receive on their way through schools dominated by the product model of knowledge in general and of writing in particular. The instruction and testing hinders rather than helps, but it fills the students with the belief in simple right answers and the conviction that memorization rather than discovery is at the heart of writing.

As long as writing is defined as a finished product to be edited and corrected, writing remains a matter of guessing about the rules that govern correctness as defined by those in power. Such a definition is perfectly enforced by the usual multiple-choice editing test.

Thus, there is a hint of the social revolutionary behind the process/product slogan. Process teaching and evaluation not only asks the teacher to teach all students to write (even those whose home dialects are different from the teacher's); it argues for complex judgments of competence based on more than the correctness of the product.

The process model sees writing as a series of overlapping activities,

We do not assign writing to students because they will turn out work that will inform us. . . . We assign writing to students so that they will learn something valuable while writing.



Edward M. White is professor of English at California State University, San Bernardino. He presented a version of this paper at AAHE's Third National Conference on Assessment in Higher Education (June 9, 1988, Chicago).

An absolutely necessary part of learning to write has to do with learning linguistic etiquette. . . . But the process model urges us to be cautious about the place of such matters in the definition of writing skill.

The very narrow definition of writing as editing puts the validity of multiple-choice usage tests in serious question despite their reliability.

all of which have to do with critical thinking and imagining: inventing and preparing, drafting, refining and rethinking, connecting, revising, and (finally) editing. The role of the teacher according to this model is very active: deviser and prompter of writing tasks, writing coach, sponsor of peer evaluations, sympathetic but questioning reader of drafts, merciless taskmaster asking for critical rethinking and revising, and (finally) evaluator.

Conservatives suspect that adopting the process model means abandoning standards, but its proponents argue that it moves standards from a rule book into a real world. This is writing seen as part of thinking, responding, evaluating, selecting and connecting ideas. Good grades become less a matter of the dialect spoken in the home (traditionally, the nearer the home dialect to the school dialect, the better the student's grade) than a measure of the thinking process of the student. The process model, by redefining writing, does not abandon standards but tries to redefine standards more substantially and more fairly. No sensible teacher using the process model abandons the requirement for a well-edited presentation copy of student work; we need to collect and grade occasional presentation copies to be sure that students can, when needed, turn out presentable products. But, the edited final draft is now just one part of the job, rather than the only thing that matters.

The social stratification issue in the process/product debate is far from trivial. The process model in part takes issue with an inescapable aspect of education in all societies: socializing students to fit into the world. This is a responsibility that no writing teacher or writing test can avoid. In a highly verbal society, an essential part of this socialization has to do with learning the conventional rules of verbal behavior.

An absolutely necessary part of learning to write has to do with learning linguistic etiquette, memorizing the way things are done in the school dialect, internalizing the writing patterns of those in power. We do not want original approaches to spelling, punctuation, or sentence structure (though we often get them); students should learn the rules on such matters and conform. But the process model urges us to be cautious about the place of such matters in the definition of writing skill.

Take spelling, for instance, the archetypal socializing study, the darling of those who in the name of "standards" want education to focus on rote drill. Spelling is important because manners are important for social mobility, and good spelling suggests good linguistic upbringing. Spelling is important not because it makes sense nor because it is good mental discipline nor because it has anything to do with thinking; spelling is important solely because people will think you are stupid if you fail to spell words the way everyone else spells words.

But education in a democratic society cannot rest content with teaching and testing conformity and good manners, useful though these traits are for the society and for the students eager to fit into that society. When Thomas Jefferson and his colleagues (five of whom founded universities) spoke of the need for public education, as they frequently did, they were not concerned about spelling or job skills or even competing with the Japanese for world markets. They were concerned with maintaining a government whose powers derived from the consent of the governed. The problem for democratic political theory is that an uninformed consent is not consent at all.

Behind all the volumes talking about school curriculum lies that simple fact: education in a democracy must prepare an informed and independent electorate. Despite the necessary conformity and socialization that must go on in school, any school, we have the particular obligation to foster critical thinking, questioning, and creativity. And writing, as defined by the process model, turns out to be the single most effective way to develop these most essential basic skills, at all levels. Any definition of

basic skills in writing that leaves out thinking is undemocratic and insubstantial.

Let me note that many other fields have also been dealing with a similar process/product conflict in definition. Reading theorists these days point out that readers are not merely passive recipients of a given text but rather active participants in the creation of the meaning of that text; reading and writing are similar meaning-making activities. And mathematical reasoning [as Charles Pine noted in the September *Bulletin*] is clearly more important than calculation and formulas that may lead to right answers. In every case, the product definition, based on rules for correctness, leads to easy test making but to a simplified curriculum, unsuited to free men and women.

The implications of the process/product issue for the testing of writing are profound. They can be examined in terms of the three kinds of measurement now in use: multiple-choice usage tests, essay tests, and portfolio assessments.

Multiple-choice usage tests. The familiar multiple-choice usage tests are, of course, based on the definition of writing as a product. The assumptions are plain: (1) There are clear right and wrong answers to problems in the editing of texts. (2) The brief fragments of writing given on the test can elicit a student's knowledge of those answers. (3) The student's ability to write (vaguely defined) will correlate highly with ability to edit correctly the prose on the test. Unfortunately for educators committed to this set of assumptions, the process model of writing disputes each of them.

The advantages of this kind of test are familiar to everyone and result from the established traditional methods of devising a reliable norm-referenced test. Once the very high development costs are covered, the test can be easily administered and scored at low cost by computer. Once an appropriate norm group has been established, the norms can be published and score comparisons can be made. The statistics are impressive, and scoring reliability is essentially perfect.

Best of all, professional educators love these tests and the data produced by them, and the tests have the resounding appearance of objectivity. (Some people even call them "objective tests," as if they had been constructed by the same computers that score them.)

These are powerful advantages—powerful enough for multiple-choice usage tests to survive and even to thrive despite the attacks of process advocates. But, if we do think of writing as a thinking and creating process, the disadvantages of the tests far outweigh their advantages. The problem is not alone that they test only a minor part of the construct "writing"; the nature of the test itself, with its right "keys" and wrong "distractors," opposes the development of critical thinking.

In more technical terms, the very narrow definition of writing as editing puts the validity of multiple-choice usage tests in serious question despite their reliability. The validity issue becomes even more serious as evidence accumulates that the supposed correlation of scores on these tests with measures of actual writing is much higher for students who call themselves "white" than for students who call themselves "minority." As long as the validity of the tests is so suspect, none of the other advantages should carry the weight they now do.

Essay tests. In response to the validity problem, we have seen major advances in essay testing over the last generation. Asking students to write on a writing test solves many, but by no means all, of the problems I have just alluded to.

Writing is still defined as a product to be measured, so the process model is still unsatisfied by an essay test. But at least such tests measure a much wider range of writing activity than mere editing of test prose. Students must demonstrate the ability to generate prose and to develop thought in prose. The measurement can respond to these parts of the writing process, as well as to editing in

Conservatives suspect that adopting the process model means abandoning standards, but its proponents argue that it moves standards from a rule book into a real world.

The teacher who at heart considers school writing as a product will serve basically as a more-or-less-negative evaluator, red-marking perceived violations of some set of rules.

Partisans of the process model are never altogether comfortable with tests that exclude revising, which for them is the most essential part of the writing process.

Of all assessment methods, portfolio assessment is the most responsive to the process model of writing. Whether it can solve the problems of reliability and cost remains an open question.

a more natural setting.

Of course, such a test is still an artificial situation for writing and still a small sample of the universe of writing, so we cannot consider the writing produced to be "real writing," as some uncritical proponents of essay testing put it.

Essay tests, nonetheless, have become more and more widely used nationally in recent years. They have become standard on college placement tests, for example, and on teacher certification skills tests, though they are often combined with multiple-choice components to increase reliability. Most such tests are carefully developed, using recognized procedures calling for test criteria and pretesting of topics; rating reliabilities and costs (always the twin problems for writing tests) have become manageable using controlled holistic scoring.

The advantages of essay tests are clear. Their increased validity, adequate reliability, and nearness to the process model of writing have earned them support from teachers and many administrators. The evidence of their relative freedom from racial bias has been impressive. Furthermore, because teachers must necessarily be involved in creating and scoring essay tests, substantial in-service training of teachers has been a useful by-product of their use.

But the disadvantages of essay testing are also obvious. Development costs are much lower than for multiple-choice usage tests, but scoring costs are higher. Although it has become more and more customary to achieve very high inter-rater score correlations, it is difficult to achieve high overall test reliability with a single essay. If test reliabilities above 0.5 are needed, evaluators must obtain and score more than one writing sample, or they must combine an appropriate multiple-choice component such as a reading test with the essay. And test development is a never-ending process.

Essay testing is always a series of compromises. Partisans of the process model are never altogether comfortable with tests that exclude revising, which for them is the most essential part of the writing process and an essential part of the construct being tested.

Portfolio assessments. In response to these problems, portfolio assessment has recently been adopted from the fine arts to writing measurement. For example, the University of Minnesota has announced that it will evaluate writing portfolios as the test for entry into freshman English; SUNY at Stony Brook has begun using portfolios to measure upper-division writing proficiency; and even the Educational Testing Service has continued to support and publicize a portfolio research project.

Of all assessment methods, portfolio assessment is the most responsive to the process model of writing. Whether it can solve the problems of reliability and cost, as essay testing has, remains an open question. At present, portfolios are an intriguing but experimental approach to writing assessment. I suspect that the approach will not work for large-scale assessment, but that's what specialists were saying about essay testing a generation ago, and they turned out to be wrong.

The measurement of writing ability should support the best goals of writing instruction. This means that the measurement as well as the teaching of basic skills must include critical thinking and creativity in the definition of these basic skills. It also means that the processes of thinking and writing, as well as their products, need to appear somehow in the measurement of writing ability. We are clearly a long way from realizing that goal, both in the testing and the teaching of writing.

A community college teacher once told me of a one-sentence essay she received from a remedial-level student. Responding to the topic "why write?" the student wrote, "They make you write so they can getcha." For this student, not only the test but all school writing was a trap. Our task and our challenge is to help such students see the opportunity that writing offers to them. ■